PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/62, 15/12, 15/32, C07K 19/00, 14/325, 14/435, A01N 63/02, A01H 5/00

A2

(11) International Publication Number: **WO 95/11305**

(43) Internati nal Publication Date: 27 April 1995 (27.04.95)

(21) International Application Number: PCT/GB94/02274

(22) International Filing Date: 18 October 1994 (18.10.94)

(30) Priority Data:
9321469.0 18 October 1993 (18.10.93) GB

(71) Applicant *(for all designated States except US)*: ZENECA LIMITED [GB/GB]; 15 Stanhope Gate, London W1Y 6LN (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: ELY, Susan [US/US]; 140 Storm Road, Groton, New York, NY 13073 (US).

(74) Agents: ROBERTS, Timothy, Wace et al.; Intellectual Property Group, Zeneca Seeds, Jealott's Hill Research Station, P.O. Box 3539, Bracknell, Berkshire RG12 6YD (GB).

(81) Designated States: AU, BB, BG, BR, BY, CA, CN, CZ, FI, GE, HU, JP, KG, KP, KR, KZ, LK, LT, LV, MD, MG, MN, NO, NZ, PL, RO, RU, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: INSECTICIDAL PROTEINS

(57) Abstract

Chimeric insecticidal proteins comprise at least part of a Bacillus thuringiensis δ-endotoxin fused to a venom-derived insecticidal protein, such as the AaHIT peptide obtainable from Androctonus australis Hector. The δ-endotoxin portion protects the venom-derived protein and delivers it to the insect gut. DNA constructs encoding such chimeric proteins may be used to express said proteins in biological organisms. Exposure of insects to the chimeric insecticidal proteins is achieved through application to plants of an insecticidal composition containing said proteins or through expression of said proteins within transgenic plants.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

INSECTICIDAL PROTEINS

The present invention relates to novel insecticidal proteins and uses thereof.

The organism Bacillus thuringiensis produces a crystal-associated protein δ-endotoxin which kills insect larvae upon ingestion. It is not, however, toxic to mammals. It is thus very useful as an agricultural insecticide, in particular against Lepidoptera, Coleoptera and Diptera. Strains of Bacillus thuringiensis have been used as agricultural insecticides for a number of decades.

B thuringiensis δ-endotoxins include those specifically insecticidal to lepidopteran larvae (such as the CryI type proteins: Hoefte and Whiteley, 1989, Microbiol. Rev., 53:242-255), those specifically insecticidal to coleopteran larvae (such as the CryIII type proteins: Krieg et al, 1893, J. Appl. Entomol., 96:500-508) and those with dual specificity for lepidoptera and coleoptera (such as the CryV protein: International patent application publication number WO90/13651; European patent application publication number 474662; Tailor et al, 1992, Molec. Micro., 6:1211-1217). Plant transformation constructs encoding such proteins have been used to produce insect resistant transgenic plants (Ely, 1993, in Bacillus thuringiensis, An Environmental Biopesticide: Theory and Practice, eds. Entwistle et al, pp 105-124).

B thuringiensis δ-endotoxins are large, fairly insoluble proteins which are produced as parasporal crystals, and which often aggregate when in

solution. The lepidopteran-active CryI type δ-endotoxins are synthesized as precursor molecules which are "activated" by trypsin-like cleavage in the larval midgut. These activated core endotoxins are then resistant to further degredation by insect gut enzymes. After ingestion, the δ-endotoxins bind to larval midgut epithelium; the αhelical N-terminal domain of the protein is then thought to penetrate the epithelial cell membrane. At this stage δ-endotoxin binding is irreversible and midgut integrity is destroyed, eventually leading to osmotic disequilibrium and cessation of feeding. Because this is a multistep process, larval death is relatively slow and may take from three to seven days.

B thuringiensis δ-endotoxins have been shown to retain insecticidal activity when in 5' translational fusion to "marker" proteins such as neomycin phosphotransferase (Hoefte, 1988, FEBS Lett., **226**:364-370) and to be correctly processed (Honee, 1990, Appl. Env. Micro., **56**:823-825).

Attempts have been made to extend the host range of B thuringiensis δ-endotoxins, making them active against a wider range of insects. For example, International patent application publication number WO9117254-A (University of California, equivalent to US patent numbers 5143905 and 5306628) describes a chimeric protein comprising a first protein segment having insecticidal activity and a second protein segment (targetting protein) capable of binding strongly to the insect gut epithelium. The insecticidal protein is preferably a δ-endotoxin (Cry protein) of Bacillus thuringiensis. The targetting protein

is a viral protein (such as gp64 glycoprotein of Autographa california Nuclear Polyhedrosis Virus) or a bacterial protein having high affinity for the lipid components of membranes. The Bacillus thuringiensis δ-endotoxins (Cry proteins) are not used as targetting proteins. WO9117254-A further describes DNA comprising a first DNA fragment encoding an insecticidal protein domain and a second DNA fragment encoding a targetting protein domain. Such chimeric proteins can be expressed in transgenic plants.

European patent application number 340948 (Mycogen Corp, equivalent to US patent number 5290914) describes a hybrid pesticide protein toxin. The hybrid toxin comprises a cytotoxic agent and a gut epithelial cell recognition portion of a Bacillus thuringiensis protein. The Bacillus thuringiensis protein portion is located at the C-terminal end of the hybrid protein. The cytotoxic agent is specifically a ribosome-inactivating enzyme (preferably derived from plant seeds) or an ADP-ribosylating enzyme (such as diphtheria toxin). Such cytotoxic agents may be highly toxic to mammals.

Various taxonomic groups use "venoms" to immobilize prey. These groups include spiders, scorpions, snakes (Quistad et al, 1992, J. Econ. Entomol., 85:33-39), mites (Tomalski and Miller, 1991, Nature, 352:82-85), marine worms (Bon et al, 1985, Neurochem. Int., 7:63-75) and others. Venoms are herein defined as biologically active secretions comprising proteinaceous components; non-proteinaceous components may also be present in such secretions. The components of many venoms

include small proteins which are insecticidal. For example, the tox34 family of related proteins is obtainable from female mites of the species Pyemotes tritici. Certain venom-derived insecticidal proteins have been shown to be highly specific for insect sub-cellular structures such as sodium channels (McCutchen et al, 1991, Bio/Technology, 9:848-852), and to be non-toxic to mice (Stewart et al, 1991, Nature, 352:85-88; Tomalski and Miller, 1991, Nature, 352:82-85). Therefore these venom-derived insecticidal proteins are potentially useful as agricultural insecticides. For example, European patent application publication number 431829 (Agracetus Inc, equivalent to US patent number 5177308) describes insecticidal proteins produced by insect predatory Arthropods, including the insect-specific neurotoxin AaHIT from the scorpion Androctonus australis Hector, the toxins BeIT1 and BeIT2 from the scorpion Buthus epeus, agatoxins from the spider Agelenopsis aperta, and a neurotoxin from the spider Sestria florentina. Such toxins may be encoded by a genetic construct for expression in various transgenic host organisms, including microorganisms and plants.

European patent application 431829 (Agracetus) also describes transgenic plants containing two separate genetic constructs, each encoding a separate insecticidal toxin of different toxicity. For example, the first construct may encode a Bacillus thuringiensis δ-endotoxin while the second construct encodes an Arthropod-derived insecticidal protein. The presence of these two different toxins in the same plant may delay development of resistance by susceptible insects.

Venoms are normally delivered by injection; they are either inactive or of low insecticidal activity if ingested (Agracetus, European Patent Application 431829), presumably because the venom's components are degraded by insect gut digestive enzymes. This severely limits the usefulness of venom-derived proteins as insecticides.

It is one object of this invention to provide an improved method to deliver a venom-derived insecticidal protein to the insect gut.

According to the present invention we provide a chimeric protein comprising an N-terminal portion fused to a C-terminal portion, the N-terminal portion being at least part of a Bacillus thuringiensis δ-endotoxin including the α-helical N-terminal domain and the gut epithelial cell recognition domain and the C-terminal portion being a venom-derived insecticidal protein.

The N-terminal portion and the C-terminal portion are both obtainable from natural sources. The B thuringiensis δ-endotoxin portion of the chimeric protein may be insecticidally active. The venom-derived insecticidal protein may be any protein or polypeptide which is normally insecticidal when injected but inactive or less active when ingested by an insect. A variety of suitable small, pharmacologically-active, insect-selective toxophores are obtainable from venoms of various taxonomic groups, including Arthropods such as scorpions and spiders, mites, snakes, marine worms and others. For example, venom-derived insecticidal proteins include the tox34 proteins

obtainable from female mites of the species Pyemotes tritici, the insect-specific neurotoxin AaIT from the scorpion Androctonus australis Hector, the toxins BeIT1 and BeIT2 from the scorpion Buthus epeus, agatoxins from the spider Agelenopsis aperta, and a neurotoxin from the spider Sestria florentina. Being insect-specific toxophores, such venom-derived insecticidal proteins have no known mammalian toxicity and may therefore be used more easily for agricultural purposes.

The two separate portions of the chimeric protein (the δ-endotoxin and the venom-derived protein) are encoded by a single recombinant gene and are bound efficiently together into one fusion protein during protein synthesis in the chosen transformed host organism. The δ-endotoxin portion is not merely acting as an insecticidal protein but is used to enfold and protect the venom-derived insecticidal protein, preventing its break-down during ingestion. Furthermore, fusion to the δ-endotoxin portion delivers the venom-derived protein to the insect gut. The gut epithelial cell recognition domain of the δ-endotoxin portion allows the chimeric protein to attach to the gut wall (such that the chimeric protein has the normal host range of the separate δ-endotoxin). The α-helical N-terminal domain then causes breakdown of the gut wall. Cleavage may then release the venom-derived insecticidal protein which may pass through the disintegrating insect gut wall. Thus the chimeric protein has an advantage over the separate venom-derived protein as its insecticidal activity will be maintained when ingested by an insect. The chimeric protein may also show

improved activity over the separate Bacillus thuringiensis δ-endotoxin as the speed of larval death is increased.

The invention further provides a DNA construct capable of expressing a chimeric protein comprising a DNA sequence encoding at least part of a Bacillus thuringiensis δ-endotoxin including the α-helical N-terminal domain and the gut epithelial cell recognition domain in 5'-translational fusion to a venom-derived insecticidal protein. The orientation of the δ-endotoxin portion with respect to the venom-derived protein is essential: the δ-endotoxin comprises the 5' end of the construct and the venom-derived protein comprises the 3' end of the construct.

The chimeric fusion protein is encoded by one DNA construct comprising a chimeric gene in the correct orientation to regulatory elements (such as promoters, ribosome-binding sites, termination sequences, introns) appropriate for the chosen transgenic host (microorganism or plant).

E coli strains carrying cloned genes for chimeric translational fusion proteins may be prepared by growing cells to stationary phase on solid nutrient media (such as L agar) prior to scraping cell growth from the agar surface, lyophilising, and freezing for storage. This material may be tested directly for insecticidal efficacy if appropriate control strains (such as the same E coli carrying only the cloning vector) are included in the insect bioassays. Alternatively, the chimeric genes may be cloned in an E coli vector that facilitates purification of

the cloned fusion protein by affinity chromatography. E coli cloning vectors and affinity purification methods and materials are commerically available (for example, those from New England Biolabs, Beverly, MA). Expression may also be carried out using known methods in eukaryotic systems such as yeast, or plants. Pure fusion proteins prepared from any of these expression systems may be used in insecticidal compositions by the addition of any desired and appropriate formulating agents. Formulating agents which may be useful include, for example, surface active agents (such as wetting agents), solid diluents, dispersing agents and UV stabilisers. If desired, solid formulations may be prepared by known methods.

The invention further provides a method to combat insects by exposing them to a chimeric protein according to the invention.

In the chimeric proteins of the invention, orally-active B thuringiensis δ-endotoxins are used as delivery molecules for venom-derived insecticidal proteins normally delivered by injection. Thus non-Bacillus thuringiensis toxophores may be effectively transported to the insect gut. The 5'-translational fusion of a B thuringiensis δ-endotoxin with a pharmacologically-active insecticidal molecule allows delivery of the latter to the insect mid-gut encased in a large, fairly insoluble, digestion-resistant protein. The Bacillus thuringiens δ-endotoxin portion of the chimeric protein will bind to and disrupt the midgut epithelium of susceptible insect larvae, allowing delivery of the venom-derived toxin to the

insect's haemocoel. The latter moiety may be released from the δ-endotoxin moiety by trypsin-like cleavage in the insect gut. This provides a faster means for achieving larval death and allows venom-derived insecticidal proteins to be delivered orally, either by conventional application or by expression in transgenic plants.

Generally, plants infested or liable to infestation by insects are treated (eg sprayed) with insecticidal compositions as described above diluted with a diluent such as water. The insecticidal agent is the chimeric fusion protein; if desired this may be applied to the plants independently of the bacterial, yeast or plant expression system that produce it.

The invention also provides a transgenic plant containing DNA encoding a chimeric protein of the invention.

The plant susceptible to insect attack may produce the chimeric fusion protein in situ. This is achieved by cloning the chimeric gene by known means, providing it with a suitable promoter (for example the CaMV35S promoter) which will cause expression of the gene in plants, and transforming the plant by known methods (such as bombardment of plant suspension cells with DNA-coated particles or Agrobacterium-mediated transformation). Any moncotyledonous or dicotyledonous plant species may be transformed with a DNA construct encoding the chimeric insecticidal protein.

SEQ ID NO 1 shows the single-stranded form of a synthetic AaHIT gene with flanking regions for

cloning. Although this sequence encodes the AaHIT insect-specific peptide toxophore from the North African scorpion Androctonus australis Hector, the codon usage differs from the naturally-occuring gene. The sequence has been altered to optimise expression in E coli and dicotyledonous plants and to introduce unique restriction sites into the gene or flanking regions. The synthetic gene may be used in DNA constructs encoding a chimeric protein according to the invention. Furthermore, Example 1 illustrates a method of creating a trypsin cleavage site within the chimeric protein to allow release of the AaHIT protein moiety in the insect gut. In fact, this is a generally-applicable method for creating a trypsin-sensitive cleavage site within a chimeric protein to allow release of a protein moiety cloned in 3'-translational fusion as part of the chimeric protein. In the process of constructing a DNA sequence encoding the chimeric protein (having a first protein moiety in translational fusion to a second protein moiety), a sequence encoding an in-frame trypsin-sensitive cleavage site is introduced between the first and second protein moieties. The second protein moiety may thus be released from the fusion by trypsin-like enzymes.

The following examples illustrate the invention with reference to the Sequence Listing in which:

SEQ ID NO 1 shows the sequence of a synthetic AaHIT gene with flanking regions for cloning;

SEQ ID NO 2 shows the amino acid sequence of a trypsin recognition site;

SEQ ID NO 3 shows the sequence of a 162-mer sense strand longmer;

SEQ ID NO 4 shows the sequence of a 162-mer antisense strand longmer;

SEQ ID NO 5 shows the sequence of a 34-mer sense strand primer;

SEQ ID NO 6 shows the sequence of a 34-mer sense strand primer;

SEQ ID NO 7 shows a fragment of the wildtype CryIA(c) gene sequence starting at base number 1984;

SEQ ID NO 8 shows a fragment of the CryIA(c) protein sequence from amino acid residues 696 to 699;

SEQ ID NO 9 shows a fragment of the PCR-modified CryIA(c) gene sequence, starting at base number 1984;

SEQ ID NO 10 shows the sequence of the fusion region of the chimeric CryIA(c)-AaHIT DNA construct;

SEQ ID NO 11 shows the sequence of the fusion region of the chimeric CryIA(c)-AaHIT protein.

EXAMPLE 1

Construction of a fully synthetic AaHIT gene

SEQ ID NO 1 shows the single-stranded form of a synthetic AaHIT gene with flanking regions for cloning. The sequence has been altered to optimise expression in E coli and dicotyledonous plants and to introduce unique restriction sites into the gene or flanking regions.

In the synthetic gene, the open reading frame (ORF) encoding the AaHIT peptide begins with the ATG codon at positions 42-44, and terminates at the TAA codon at positions 256-258. The ORF contains unique recognition sites for the following

restriction endonucleases: PmeI (bases 182-189), XhoI (bases 205-210), and PacI (bases 251-258). The double-stranded version of SEQ ID NO 1 is designed with the following engineering features to be used, for example, when the sequence is cloned into the E coli vector pUC19 at the EcoR1 and HindIII sites of the pUC19 polylinker after the vector has had the single NdeI site removed by known methods (such as digestion, Klenow polymerase fill-in and religation).

(a) Method of creating a blunt-ended DNA fragment beginning at the ORF start codon

Digestion with BspM1 (recognition site at bases 33-38; cleavage at bases 42 on the sense strand and 46 on the antisense strand) followed by filling in of the single-stranded DNA ends using Klenow polymerase in the presence of dNTPs produces a blunt-end DNA fragment beginning with base 43, the first base in the AaHIT ORF. This feature allows exact fusion products to be formed when used in conjunction with appropriate restriction and modifying enzymes to manipulate the 3' end of the AaHIT gene.

(b) Method of creating a trypsin cleavage site to release the venom-derived insecticidal protein moiety

The cloned AaHIT gene is digested with NdeI at the now-unique NdeI site (bases 40-45) followed by filling in of the single-stranded ends with Klenow polymerase in the presence of dNTPs. This is followed by digestion with the blunt-cutting enzyme NruI at the unique NruI site (bases 27-32). Re-ligation results in an altered sequence in which the pair of bases, 28 and 29, are adjacent base 52,

thus producing an in-frame Arg codon (CGT). The amino acid sequence with the codon begining with the C at position 19 is now in-frame relative to the AaHIT ORF and reads "Arg, Gly, Arg, Arg, Met" (SEQ ID NO 2), with the Methionine being the start codon for AaHIT. This amino acid sequence constitutes a trypsin recognition site with cleavage between Arg and Met, and allows release of the separate AaHIT peptide from the chimeric protein by trypsin-like enzymes.

(c) Method to effect in-frame fusions of genes encoding the portions of the chimeric protein

Cleavage at the unique EcoN1 site (bases 13-23) is designed to facilitate in-frame fusions to the C-terminal portion of any B thuringiensis δ-endotoxin protein. It may also be useful to incorportate an in-frame polyproline hinge at the EcoN1 site between the δ-endotoxin and the AaHIT portions of the chimeric gene. Polyproline hinges may be of various lengths. Such hinges may facilitate correct folding of the AaHIT moiety and may provide better trypsin-mediated cleavage and release of the AaHIT moiety in the larval gut.

(d) Construction of a synthetic AaHIT gene with flanking DNA sequences using the polymerase chain reaction (PCR)

Synthetic long deoxyoligonucleotides (longmers) and short primer oligonucleotides (primers) are constructed as follows:

a 162-mer sense longmer (SEQ ID NO 3);
a 162-mer antisense strand longmer (SEQ ID NO 4);
a 34-mer sense strand primer (SEQ ID NO 5);
a 34-mer sense strand primer (SEQ ID NO 6).

Longmers were synthesized on an Applied Biosystem

model 380B DNA synthesizer on a 40 nmole synthesis scale, under conditions in which the coupling efficiency was at least 98%. In the case of the longer synthetic products (the two antiparallel 164-mers), trityis were monitored visually throughout the synthesis, and the synthesizer was paused during the synthesis to replenish the phosphoramidites. Absorbance at A_{280nm} was used to estimate oligomer concentrations. Longmers and primers were deprotected at 60°C for 39 or 15 hours respectively, prior to freeze-drying and storage at -20°C. Further purification of longmers was not required. Each longmer synthesis results in a mixture of molecules of varying lengths, all of which have identical 3′ ends. The full length 164-mer sense and antisense oligomers possess a 5o base pair overlap region.

Synthetic longmers were resuspended in sterile, deionized, glass-distilled water to a final concentration of 2 μg/μl, aliquoted and stored at -20°C. Synthetic primers were resuspented in sterile, deionized, glass-distilled water to a final concentration of 0.1μg/μl, aliquoted and stored at -20°C.

Polymerase chain reactions were done in two "rounds". Round 1 provided core gene synthesis(that is, extension of longmers in the absence of the flanking primers). Round 1 consisted of a 5′ denaturation period at 94°C prior to addition of the AmpliTaq™ polymerase and 10 cycles of:

melt 1′ at 94°C
anneal 2′ at 65°C
extend 1′ at 72°C.

Round 2 provided core gene amplification in the

presence of the flanking primers. For Round 2, 1/25 of Round 1 reaction was added to a new reaction mix prior to 30 cycles of:

melt 1' at 94°C
anneal 2' at 65°C
extend 1' at 72°C.

Reaction mixes were composed as follows:

Round 1 - Core Gene Extension (final concentrations, 50µl reactions):

50mM KCl
10mM Tris ph 8.3 at room temperature
2.5mM $MgCl_2$
0.25mM each dNTP
2.5U AmpliTaq™
4µg each longmer (sense and antisense)

Round 2 - Core Gene Amplification (final concentrations, 50µl reactions):

50mM KCl
10mM Tris ph 8.3 at room temperature
2.5mM $MgCl_2$
0.25mM each dNTP
2.5U AmpliTaq™
1/25 volume Round 1 reaction mix
200ng each flanking primer

PCR products of the correct size (274 base pairs; for example, corresponding to the double-stranded version of the sequence presented in SEQ ID NO1) were observed by electrophoresis on 2% agarose gels. The synthetic PCR product was cloned using known methods.

EXAMPLE 2

Constructi n of a chimeric DNA c nstruct encoding a CryIA(c)-AaHIT insecticidal fusion protein

A suitable CryIA(c) gene fragment may be isolated as a 6.6 kilobase HindIII fragment from the DNA of recombinant E coli strain NCIB 40211 obtained from the National Collection of Industrial and Marine Bacteria (23 St Machar Drive, Aberdeen, Scotland AB2 1RY). The cryIA(c) gene is then engineered as an in-frame 5'-translational fusion to the amino terminus of the AaHIT gene at any point downstream of the natural trypsin-sensitive site between amino acid residues 622 and 623 of CryIA(c). DNA sequences added between the CryIA(c)- and AaHIT encoding portions may include those encoding a polyproline hinge and may also include those which add a trypsin-sensitive site. One method for constructing this fusion involves production of a synthetic C-terminal region for cryIA(c) by known polymerase chain reaction (PCR) methods such that an insecticidal gene product is encoded by the cryIA(c)-derived sequence. A variety of possible lengths exist for the cryIA(c)-derived moiety.

As an example, this method was used to alter the CryIA(c) coding sequence for amino acids 696 to 698 such that the natural protein sequence was maintained but an EcoN1 site was created in the DNA sequence as shown below. The following sequence is from the wildtype CryIA(c) starting at base number

1984 (SEQ ID NO 7) and encodes amino acid residues 696 to 699 (SEQ ID NO 8):

```
5'...C C A G A A C G T G G G...3'
  ...G G T C T T G C A C C C ...
     Pro   Glu   Arg   Gly
```

The following sequence is from the PCR-modified CryIA(c) starting at base number 1984 (SEQ ID NO 9) and encoding amino acid residues 696 to 699 (SEQ ID NO 8). The PCR-altered bases are shown in bold and marked by the symbol °. The EcoN1 site is underlined and the cleavage sites are indicated by the symbols ↑ and ↓.

```
                ↓
5'...C C T° G A A C G A° G G G...3'
  ...G G A  C T T G C T  C C C ...
                ↑
     Pro    Glu   Arg    Gly
```

Digestion of the altered cryIA(c) gene with EcoN1 then allowed construction of a chimeric gene by altering the AaHIT gene as described above to create a site encoding a trypsin-sensitive peptide bond, prior to digestion with EcoN1 and ligation to the digested cryIA(c) gene by known methods. This results in a CryIA(c)-AaHIT chimeric protein, the fusion region of which is shown below. The following sequence is the sense strand (SEQ ID NO 10) encoding the fusion region of the chimeric CryIA(c)-AaHIT protein (SEQ ID NO 11). Sequences derived from the B thuringiensis CryIA(c) gene are shown in bold (five bases and two amino acids at the 5' end). The remaining sequence derives from DNA SEQ ID NO 1, after manipulation as described above. The asterisk (*) indicates the

trypsin-sensitive cleavage site. Met is the start codon for the AaHIT gene.

5'...C C T G A A C G A G G A C G T C G T A T G...3'
...**Pro Glu Arg Gly Arg Arg*** Met...

EXAMPLE 3

Construction of a chimeric DNA construct encoding a CryV-AaHIT insecticidal fusion protein

A suitable cryV gene may be isolated the DNA of recombinant E coli strain NCIB 40278 or from the DNA of B thuringiensis strain NCIB 40091 obtained from the National Collections of Industrial and Marine Bacteria (23 St Machar Drive, Aberdeen, Scotland AB2 1RY). The cryV gene is then engineered as an in-frame 5'-translational fusion to the amino terminus of the AaHIT gene at any point downstream of CryV amino residue 648. DNA sequences added between the CryV- and AaHIT-encoding portions may include those encoding a polyproline hinge and may also include those which add a trypsin-sensitive site. One method for constructing this fusion involves production of a synthetic C-terminal region for cryV by known polymerase chain reaction (PCR) methods such that an insecticidal gene product is encoded by the cryV-derived sequence. A variety of possible lengths exist for the cryV-derived moiety. This provides an in-frame cloning site anywhere from CryV amino acid residue 648 to residue 720, and an encoded trypsin cleavage site just upstream of the AaHIT initiation codon.

EXAMPLE 4

Construction of a chimeric DNA construct encoding a CryIIIA-AaHIT insecticidal fusion protein

A suitable cryIIIA gene may be isolated as a 3.0 kilobase DNA frament of HindIII-digested genomic DNA from B thuringiensis strain NCIB 40023 obtained from the National Collections of Industrial and Marine Bacteria (23 St Machar Drive, Aberdeen, Scotland AB2 1RY). The cryIIIA gene is then engineered as an in-frame 5'-translational fusion to the amino terminus of the AaHIT. DNA sequences added between the CryIIIA- and AaHIT-encoding portions may include those encoding a polyproline hinge and may also include those which add a trypsin-sensitive site. One method for constructing this fusion involves production of a synthetic C-terminal region for cryIIIA by known polymerase chain reaction (PCR) methods such that an insecticidal gene product is encoded by the cryIIIA-derived sequence. A variety of possible lengths exist for the cryIIIA-derived moiety. This provides an in-frame cloning site in the CryIIIA coding sequence and an encoded trypsin cleavage site just upstream of the AaHIT initiation codon.

EXAMPLE 4

Construction of a chimeric DNA construct encoding a CryIIIA-AaHIT insecticidal fusion protein

A suitable cryIIIA gene may be isolated as a 3.0 kilobase DNA frament of HindIII-digested genomic DNA from B thuringiensis strain NCIB 40023 obtained from the National Collections of Industrial and Marine Bacteria (23 St Machar Drive, Aberdeen, Scotland AB2 1RY). The cryIIIA gene is then engineered as an in-frame 5'-translational fusion to the amino terminus of the AaHIT. DNA sequences added between the CryIIIA- and AaHIT-encoding portions may include those encoding a polyproline hinge and may also include those which add a trypsin-sensitive site. One method for constructing this fusion involves production of a synthetic C-terminal region for cryIIIA by known polymerase chain reaction (PCR) methods such that an insecticidal gene product is encoded by the cryIIIA-derived sequence. A variety of possible lengths exist for the cryIIIA-derived moiety. This provides an in-frame cloning site in the CryIIIA coding sequence and an encoded trypsin cleavage site just upstream of the AaHIT initiation codon.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: ZENECA LIMITED
(B) STREET: 15 STANHOPE GATE
(C) CITY: LONDON
(D) STATE: ENGLAND
(E) COUNTRY: UNITED KINGDOM
(F) POSTAL CODE (ZIP): W1Y 6LN

(ii) TITLE OF INVENTION: INSECTICIDAL PROTEINS

(iii) NUMBER OF SEQUENCES: 11

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: GB 9321469.0
(B) FILING DATE: 18-OCT-1993

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 274 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGAATTCCCG GGCCTGAACG AGGACGTCGC GAACCTGCGC ATATGAAAAA AAACGGTTAC 60

GCAGTTGATT CTTCTGGTAA AGCGCCAGAA TGTTTGTTGT CTAACTACTG TAACAACCAA 120

TGTACCAAAG TTCACTACGC TGATAAAGGC TACTGTTGTT TGTTGTCTTG TTACTGTTTT 180

GGTTTAAACG ATGATAAAAA AGTTCTCGAG ATTTCTGATA CCCGTAAATC TTACTGTGAT 240

ACCACCATTA TTAATTAAGA GCTCAAGCTT GGGC 274

AGTAGCCTTT ATCAGCGTAG TGAACTTTGG TACATTGGTT GT 162

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGAATTCCCG GGCCTGAACG AGGACGTCGC GAAC 34

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCCCAAGCTT GAGCTCTTAA TTAATAATGG TGGT 34

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CCAGAACGTG GG 12

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Pro Glu Arg Gly
1

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CCTGAACGAG GG 12

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CCTGAACGAG GACGTCGTAT G 21

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Pro Glu Arg Gly Arg Arg Met
1 5

CLAIMS

1. A chimeric protein comprising an N-terminal portion fused to a C-terminal portion, the N-terminal portion being at least part of a Bacillus thuringiensis δ-endotoxin including the α-helical N-terminal domain and the gut epithelial cell recognition domain, and the C-terminal portion being a venom-derived insecticidal protein.

2. A chimeric protein as claimed in claim 1 in which the venom-derived insecticidal protein is an AaHIT peptide obtainable from Androctonus australis Hector.

3. A DNA construct capable of expressing a chimeric protein comprising a DNA sequence encoding at least part of a Bacillus thuringiensis δ-endotoxin including the α-helical N-terminal domain and the gut epithelial cell recognition portion in 5'-translational fusion to a venom-derived insecticidal protein.

4. A DNA construct as claimed in claim 3 in which the venom-derived insecticidal protein is an AaHIT peptide obtainable from Androctonus australis Hector.

5. A DNA construct as claimed in claim 4 in which the sequence encoding the venom-derived insecticidal protein is SEQ ID NO 1.

6. A method to combat insects comprising exposure of the insects to a chimeric protein as claimed in claim 1.

7. A transgenic plant comprising a DNA construct as claimed in claim 3.

8. A method to create a trypsin-sensitive cleavage site within a chimeric protein, said chimeric protein having a first protein moiety in translational fusion to a second protein moiety, which comprises cloning a DNA sequence encoding the chimeric protein and introducing a region encoding an in-frame trypsin-sensitive cleavage site between the first and second protein moieties such that, when the encoded chimeric protein is expressed, the second protein moiety is releasable from the fusion by trypsin-like enzymes.

Chimeric gene for the transformation of plants.

Patent Number:	EP0508909
Publication date:	1992-10-14
Inventor(s):	LEBRUN MICHEL (FR); LEROUX BERNARD (FR); SAILLAND ALAIN (FR)
Applicant(s)::	RHONE POULENC AGROCHIMIE (FR)
Requested Patent:	EP0508909, B1
Application Number:	EP19920420066 19920304
Priority Number(s):	FR19910002872 19910305
IPC Classification:	A01H5/00 ; C12N1/21 ; C12N5/10 ; C12N15/54 ; C12N15/62 ; C12N15/82
EC Classification:	C12N15/82C8B4, C12N15/62
Equivalents:	AU1144292, AU652610, BR9200790, CA2061636, DE69226466D, DE69226466T, ES2118802T, FR2673643, IL101115, JP5095789, KR233191, MX9200915, ZA9201645

Abstract

1) Chimeric gene for conferring on plants increased tolerance to a herbicide against EPSPS. 2) It comprises, in the direction of transcription, a promoter region, a transit peptide region, a sequence encoding tolerance to glyphosate and a polyadenylation signal region, characterised in that the transit peptide region comprises, in the direction of translation, at least one transit peptide from a plant gene encoding a plastid-localised enzyme, a portion of the sequence of the mature N-terminal portion of a plant gene encoding a plastid-localised enzyme, then a second transit peptide from a plant gene encoding a plastid-localised enzyme. 3) Production of glyphosate-tolerant plants.